CHAPTER 1: THE CHILDREN OF THE NLSY79

Introduction

This data users guide provides substantive and technical information about the NLSY79 Child and Young Adult survey data. The document begins with a discussion of the child and mother samples, their constraints, and how the samples have changed over time. This first chapter provides an overview of the NLSY79 child and young adult surveys, focusing on issues that are relevant to children of all ages. Chapter 2 discusses the Child data collection and the assessments used to measure the development and environments of the sample children over time. Chapter 3 focuses on the Young Adult data collection, including information about accompanying geocode files. Chapter 4 offers some potential research applications that use the longitudinal and intercohort aspects of the data. Chapter 5 lists specific information on how to access and use the data files and documentation.

This document should be used in conjunction with a variety of other materials including selected NLSY79 Child Handbooks, the *NLS Handbook*, the *NLSY79 User's Guide,* and the questionnaires used in the field to collect the data for children, young adults and the main Youth respondents. For in-depth information about constructed variables on the Child file that are drawn from the mother's record (in the main NLSY79 files), users should consult the *NLSY79 User's Guide*. Information on how to obtain documentation can be found in Chapter 5.

Overview of the NLSY79

The National Longitudinal Survey of Youth 1979 cohort (NLSY79) is a multi-purpose panel survey that originally included a nationally representative sample of 12,686 men and women who were all 14 to 21 years of age on December 31, 1978. Annual interviews have been completed with most of these respondents since 1979, with a shift to a biennial interview mode after 1994. This report primarily focuses on the interviews with the 6417 children of the women in the 2000 survey. As of the 2000 interview round, all the women had attained the ages of 35 to 43. The children of these female respondents as of 2000 are mostly below the age of 25 and are estimated to represent about 90 percent of all the children ever to be born to this cohort of women.

Sponsored by the U.S. Department of Labor, the NLSY79 contains extensive information about the employment, education, training, and family experiences of the respondents. The survey originally included substantial over-samples of African-American,

Hispanic, economically disadvantaged white, and military youth. Reflecting budget constraints, the latter two over-samples have been largely deleted from the sample. The remaining sample, however, retains its national representation. With appropriate weights, the NLSY79 may be considered as representative of the living members of a national sample of men and women who were 14 to 21 years of age on December 31, 1978. With appropriate weights, the children of the female respondents in this sample, the focus of this users guide, may be considered as a representative sample of children who have been born to this national sample of women. Readers seeking more detail about the main Youth sample of men and women are referred to the current edition of the *NLS Handbook* and its associated *NLSY79 User's Guide*.

The NLSY79 Child Surveys

Starting in 1986, the children of NLSY79 female respondents have been assessed every two years**.** The assessments measure cognitive ability, temperament, motor and social development, behavior problems, and self-competence of the children as well as the quality of their home environment. Since 1988, children age 10 and over have completed personal interviews that have asked about a wide range of their schooling, family, peer-related and other attitudes and behaviors. Since 1994, more extensive separate interviews/data collection efforts have been completed for the older children of the NLSY79 female respondents. This data collection, subject to constraints noted in Chapter 3, has focused on NLSY79 "children" age 15 and over as of the end of the relevant calendar year (December 31, 2000 for the year 2000 interview round). In this guide, when these older children are the focus, they will be referred to as "Young Adults." When children below age 15 are being referenced, they will be termed "younger children." In general, references to the full child cohort will be termed "children."

Prospective users should use this document in conjunction with a variety of other documents including the *NLS Handbook* and *NLSY79 User's Guide*, the *2000 NLSY79 Child Assessments: Selected Tables*, and relevant child and young adult questionnaires. All of the above materials are available either on line or from the NLS User Services Office at CHRR. Details on documentation can be found in Chapter 5 of this guide.

Substantive questions regarding the data collection and assessments for the younger children should be addressed to Paula Baker at 614-442-7375 (baker.21@osu.edu). Questions

regarding the Young Adult component of the study may be addressed to Canada Keck at 614-442-7377 (keck.2@osu.edu). General questions relating to survey content or the utility of the data set for specific research topics may be addressed to Frank Mott at 614-442-7378 (mott.1@osu.edu). Questions relating to the availability and cost of public use materials should be addressed to the NLS User Services Office at the Center for Human Resource Research, 614-442-7300 (usersvc@postoffice.chrr.ohio-state.edu). The Child and Young Adult data are also available as a free download online from <http://www.bls.gov/nls>. User comments regarding any aspect of this survey including suggestions for additions or deletions are welcomed.

The Child Samples

The NLSY79 child sample is comprised of all children born to NLSY79 mothers. With the primary support from the National Institute of Child Health and Human Development (NICHD), the children of the NLSY79 mothers have been independently followed and interviewed in various ways starting in 1986. These children have been interviewed and assessed biennially since that date. Since 1988, all of the children age ten and over have completed fairly extensive self-report questionnaires. The content and scope of the child interviews and assessments are discussed in Chapter 2. Starting in 1994, children who have reached the age of 15 by the end of the survey year are no longer assessed but instead complete personal interviews akin to those given to their mothers during late adolescence and into adulthood. Chapter 3 focuses on the content of the questionnaires administered to these Young Adult children.

The Child & Young Adult Data Collection

Interviews with the NLSY79 younger children are typically conducted in the respondent's home by experienced, specially trained field staff. Reports are obtained from the children and their mothers and by interviewers trained to directly assess each child and to provide evaluations of their home environment. Interviews with all children through 1992 were conducted primarily in person using paper and pencil. Beginning in 1994, the primary Young Adult and younger child instruments and assessments were administered using computer assisted personal interviewing (CAPI). From 1994 to 1998, the primary mode of

data collection for the Young Adults was in-person interview. In 2000, the primary interview mode for the Young Adults shifted to telephone rather then in-home visits. The field period of the interviews has largely coincided with the main interview field period, although Young Adult telephone interviewing in 2000 began approximately six weeks prior to the main and younger child interviews. Further detail about the survey interviewing procedures for the Child and Young Adult surveys may be found in Chapters 2 and 3 respectively.

Sample Sizes: Who Was Interviewed in 2000?

NLSY79 mothers and children. As of 2000, a total of 11,205 children have been identified as having been born to the original 6,283 NLSY79 female respondents, mostly during the years that they have been interviewed. A modest number were born prior to the 1979 first interview round. Obviously, an unknown number of additional children have been born to women subsequent to their having attrited from the sample.

In 2000, of the 4,113 women interviewed, 3,425 were mothers who reported a total of 8,323 children (see Table 1.1). When appropriate weights are applied, NLSY79 women have had on average about 1.82 children, which represents an estimated 90 percent of their ultimate childbearing. A very large proportion of the childbearing for this cohort is now behind them although caution is still advised when generalizing from any selected portion of this child cohort.

Table 1.1. NLSY79 Mother and Child Samples: 1986-2000 Surveys

Sample Groups	1979	1986	1988	1990	1992	1994	1996	1998	2000
NLSY79 Females									
Interviewed	6,283	5,418	5,312	4,510	4,535	4,480	4,361	4,299	4,113
NLSY79 Mothers:									
Interviewed		2,922	3,346	3,088	3,325	3,464	3,489	3,533	3,425
Interviewed; Children also interviewed		2,774	3,196	2,772	2,964	3,212	3,228	3,221	2,934
NLSY79 Children:									
Born to interviewed mothers		5,255	6,543	6,427	7,255	7,862	8,123	8,395	8,323
Children not Young Adults						6,622	6,010	5,343	4,438
Young Adults						1,240	2,113	3,052	3,885
Interviewed[1]		4,971	6,266	5,803	6,509	7,089	7,103	7,067[2]	6,417
Children not Young Adults						6,109	5,431	4,924	3,392
Young Adults						980	1,672	2,143	3,025

NOTE: Sample sizes for all child surveys exclude the 441 female members of the military subsample dropped from interviewing in 1985 and their children. In addition, sample sizes for 1990 and later exclude female members of the economically disadvantaged nonblack/non-Hispanic subsample whose children were not eligible for assessment. The exclusion of this subsample after 1990 accounts for much of the drop in sample size between 1988 and 1990.

[1] An interview was considered "complete" if an interviewer was able to directly assess a child or obtain a maternal report of the child's background, health, or assessment information as recorded in either the child or mother survey instrument.

[2] This total includes 37 children who were assessed and interviewed, but whose mothers were *not* interviewed.

Of the 2,170 NLSY79 women not interviewed in the 2000 survey round, 441 were members of the military over sample that was dropped after 1984, 890 were from the economically disadvantaged white over sample that was dropped in 1990, and 105 are identified as deceased. Excluding these subsets of respondents means that effective attrition for those who would otherwise be eligible for interview is about 15 percent.

Children & young adults eligible for interview. In all but the first round of the NLSY79 child data collection, all children born to NLSY79 women are generally eligible to be interviewed, subject to the following residential limitation. Children who are part of the younger children sample (age 0-14) must reside at least part or full time with the NLSY79 mother respondent. Young Adult children, with at least one record in the child interview history, are eligible for interview regardless of their residence status. In 1998, the Young Adult interview was limited to youth between the ages of 15 and 20. In 2000 the criteria were restricted (for this survey round only) to exclude from eligibility a random sample of about 38 percent of the children and young adults from the black and Hispanic over samples. This

restriction means that while the full set of oversample mothers was contacted in 2000, only about 60 percent of their children were part of the fielded sample targeted for interview.

Changes in the child samples. The age distribution in Table 1.2 underscores the fact that many NLSY79 women are reaching the end of their childbearing years. There has been a gradual change in the mix of the child population from being predominantly a younger child group towards being older, or more of a young adult population. As of the 2000 survey round, the overall child sample is almost equally divided between children under the age of 15 (the "younger children") and children aged 15 and over (the "young adults"). Since very few children remain to be born in forthcoming NLSY79 survey rounds, we anticipate a rapid transition towards an even older child population, with increasing numbers in their 20s and even early 30s. Paralleling this shift, the younger component of the overall sample will be increasingly from middle class households, as they will have been born to women at older ages. This trend implies that within-sample analyses comparing children at different ages need to done cautiously.

Table 1.2. Age of Child by Age of Mother at Birth of Child: NLSY79 Children and Young Adults Interviewed in 2000

Child Age	Age of Mother at Birth of Child																											Total
	≤14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40+	
<1																					1	21	26	29	13	9	10	109
1																					18	24	22	18	14	3	14	113
2																			3	20	31	36	28	20	13	8	4	163
3																		1	27	33	24	27	23	12	5	5		157
4																	2	26	39	25	29	19	11	14	5			170
5																1	29	27	40	31	28	12	13	12	2			195
6															2	25	35	34	27	25	20	11	16	1				196
7														2	39	34	41	54	35	22	19	22	2					270
8													2	36	42	41	35	33	22	6	12	4						233
9												1	38	49	54	39	38	39	24	23	4							309
10											2	35	48	47	54	42	34	34	28	3								327
11										1	29	43	50	48	48	38	22	30	7									316
12										24	42	56	54	50	32	23	23	9										313
13								3	27	51	45	36	47	35	37	33	4											318
14 C								13	24	27	20	29	23	31	25	11												203
14 Y							4	31	39	28	30	28	25	22	14													221
15						5	34	61	45	57	45	37	52	34	8													378
16					2	34	49	40	51	45	46	43	37	6														353
17				3	30	43	50	56	42	38	48	36	3															349
18				20	29	41	47	50	50	38	38	6																319
19			10	28	40	52	47	29	27	39	3																	275
20		5	22	36	44	48	53	52	43	7																		310
21	2	8	20	28	39	41	43	53	8																			242
22	5	11	25	34	34	38	41	5																				193
23	3	8	31	25	31	40	4																					142
24	1	13	32	38	36	7																						127
25	5	8	25	23	4																							65
26	3	8	11	7																								29
27	2	12	4																									18
28	2	1																										3
29	1																											1
Total	**24**	**74**	**180**	**242**	**289**	**349**	**372**	**393**	**356**	**355**	**348**	**350**	**379**	**360**	**355**	**287**	**263**	**287**	**252**	**188**	**186**	**176**	**141**	**106**	**52**	**25**	**28**	**6417**

Note: Child age for children under 15 is measured as of mother's interview date (see C00047.45) and may differ from age at date of child assessment. For two children whose mothers were not interviewed in 2000, age at Child Supplement interview date (C00070.47) was used. Age of children 15 and over (young adults) is measured at their young adult date of interview (see Y11924). Age of mother at birth of child is found in the variable C00070. in the CHILD BACKGROUND area of interest of the child data file.

In the context of the child interviews, "interviewed" for children under age 15 means that some information was obtained from either the mother or child in that survey year. The content of the child data collection and assessment information is described in some detail in Chapter 2. From the perspective of the young adult sample, a completion is defined as a case in which at least a major part of the young adult CAPI interview was completed. A series of flags in the data file (in the "area of interest" called CHILD BACKGROUND) indicate interview and assessment status for younger children and young adults.

Child interviews. Table 1.2 provides information on the number of younger children and young adults who completed the year 2000 interview. There are relatively large numbers of children at all ages below 23. Whereas at one time a large proportion of the children had been born to adolescent mothers, in 2000 fully 82 percent of all the children had been born to women age 20 and over. For the older, young adult children, the percent born to adolescent mothers is 38 percent. As noted earlier, the 2000 data collection includes all children born to these women, excluding a random component of about 38 percent of the children of the black and Hispanic over-samples. The 1998 data collection was limited to all eligible children under the age of 21 as of the date of interview. Earlier rounds included all age eligible children, subject to the fact that women in the military over-sample were dropped from the interview process after the 1984 survey round and the economically disadvantaged white over-sample with the 1990 data collection round. The sample nonetheless retains sufficient numbers of children from these categories to maintain its full national representation. Appropriate weights are available in each year to adjust the un-weighted sample cases for the minority over-samples and year-to-year sample attrition.

Sample limitations. Table 1.3 suggests one other caveat for studies that focus on the consequences of earlier, adolescent childbearing for this cohort of women. A modest proportion of the children (677 of the 6,417 interviewed in 2000) was born prior to the first NLSY79 interview round. If essential explanatory inputs for analysis include pre-1979 points (e.g., employment status in 1977 or early paternal presence in the home), sample size may be temporally constrained because of this left-censoring problem—the unavailability of some data elements for the pre-survey period. All such cases fall in the upper young adult ages, and could affect analyses for young adult children who are 22 or older in 2000.

Table 1.3. Child's Birth Year by Age of Mother at Child's Birth: NLSY79 Children and Young Adults Interviewed in 2000

Child Birth Year	Age of Mother at Birth of Child																											Total
	≤14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40-42	
1970	1																											1
1971	2																											2
1972	1	5																										6
1973	4	12	7																									23
1974	2	9	17	13																								41
1975	3	12	26	37	18																							96
1976	2	6	38	26	38	20																						130
1977	4	9	28	30	25	45	23																					164
1978	3	12	19	35	37	40	40	28																				214
1979	2	7	26	32	52	44	52	56	29																			300
1980		2	16	37	42	50	47	39	36	28																		297
1981			3	20	30	57	58	42	39	40	17																	306
1982				12	38	38	43	55	47	39	48	26																346
1983					9	43	50	50	49	36	51	42	22															352
1984						12	39	52	48	57	44	34	53	25														364
1985							20	55	57	52	50	48	42	37	22													383
1986								16	37	50	39	41	36	46	39	24												328
1987									14	39	39	44	55	42	32	30	13											308
1988										14	45	52	50	43	34	26	25	19										308
1989											15	51	55	56	61	51	23	35	17									364
1990												12	50	46	47	38	33	27	26	12								291
1991													16	46	63	31	42	37	24	15	9							283
1992														19	36	42	39	50	34	15	16	8						259
1993															21	30	38	50	27	21	18	25	3					233
1994																15	34	26	35	28	26	9	22	5				200
1995																	16	25	31	26	24	15	9	16	3			165
1996																		18	37	28	34	19	21	12	5	4		178
1997																			21	33	25	36	23	14	7	7	2	168
1998																				10	22	30	21	20	19	3	7	132
1999																					12	25	26	23	9	6	11	112
2000																						9	16	16	9	5	8	63
Total	24	74	180	242	289	349	372	393	356	355	348	350	379	360	355	287	263	287	252	188	186	176	141	106	52	25	28	6,417

The increasing heterogeneity of the child sample may also be noted in other ways from Table 1.2. For example, analyses focusing on children in the adolescent age range of ten to seventeen can now include a substantial number of children who have been born to mothers who were between the ages of 20 and 30 at the birth of the child. Having noted this increase in sample heterogeneity over the years, the user should remain mindful that the oldest and youngest children in the sample are likely to come from families that differ in their socio-economic characteristics. However, it is also becoming increasingly reasonable to generalize from the NLSY79 sample of children to broader representations of selected U.S. child populations; overall, it is worth reiterating that as of this date, the cohort of women have completed about 90 percent of their childbearing.

Selecting an Age Variable: A Cautionary Note

The NLSY79 child and young adult data files include several distinct child age variables at each survey point. The most appropriate age variable to use may depend on one's research objectives. There are four relevant age variables for younger children. In most but not all instances, the values for these age variables will be the same. Also, all of these age variables are the age of the child in months, so users who need a variable in which the unit is in "years" will need to do a simple conversion. One primary child age variable references the age of the child as of the mother's interview date. One other created age variable is linked to the date the mother supplement was administered. This variable, termed "Child age at mother supplement date," is appropriate to use when one's research utilizes a mother supplement assessment. Similarly, there is a child assessment age based on the date the interviewer-administered assessments were given. Finally, for younger children there is a PPVT age, which references the date that the child completed the PPVT. All of these issues are clarified further in Chapter 2.

Young adult age variables. There are two young adult age variables most appropriate for use. In contrast with the variables for the younger children, these are computed as age-in-years. One references the young adult's age at his or her interview date, and is the variable that most users would access. The second is an age that references the last day of the calendar year. This variable is included because December 31 of a given reference year defines the eligibility of a child for inclusion in the young adult sample. Starting in

1994, a child must be 15 or older as of that date to be included in the young adult sample. This is the reason that the tables in this report that provide child age as of an interview date split the 14-year-old group between a child and young adult component. The age at end of year variable can also be useful for defining a sample according to an unchanging age cohort as of any interview year. Interview dates are rarely exactly two years apart; an individual who is age 20 at one interview point, while typically 18 at the preceding interview, might possibly have been 19 or 17. Users should also be aware that over the interview years occasional revisions have been made to a child's date of birth that was found to be in error. However, the questions and assessments administered to a child were contingent on the child's age as specified at the time. For a variety of reasons, no attempt has been made to alter the historical age record when a date of birth was revised. Thus, if one is using age-sensitive information from prior interview points, two options are often possible; one may recreate an age variable based on the most recent date of birth of child in conjunction with the interview date in that year. The second option is to use the set of age variables from that year, a variable that will be consistent with all the other information gathered from the child in that year. In most instances, this later option is probably the preferred solution. Finally, in almost no instance would it be appropriate to simply decrement a child's age by the number of years between the most recent (in this case, 2000) interview and the interview round of interest. This can often lead to incorrect estimates for the reason noted above. Additional information about some of these issues as well as the file location of the appropriate age variables is discussed further in chapters 2 and 3.

Sample Changes over Time

The increasing heterogeneity of the child sample over time may be noted from Table 1.4. This table summarizes the age mix as well as the race/ethnic mix of the child sample as it moves forward in time from 1986, the first year of the child interviews, to 2000. Over time, there is a gradual transition towards an older average age at interview. Notwithstanding this increase in age, the overall sample has actually changed very little over time in its racial and ethnic makeup. There has been some counterbalancing between higher minority birthrates and the reality that a higher proportion of the births in recent years are to older, white women. The reader may also note from Table 1.4 that sample size variations over the years have been

considerable, largely reflecting the variations in data collection already noted; the decline from 1988 to 1990 largely reflected the removal of the economically disadvantaged white over-sample. The slight decline from 1996 to 1998 was related to the capping of interviews in that year only at age 20; and the decline from 1998 to 2000 reflected the one time exclusion of a part of the black and Hispanic over-samples. In the 2002 interview round, there are no age or other sample exclusions.

Table 1.4. NLSY79 Child Sample Sizes by Age and Race/Ethnicity

	1986	1988	1990	1992	1994	1996	1998	2000
Total Interviews	4,971	6,266	5,803	6,509	7,089	7,103	7,067	6,417
By Age								
Birth to 9 Years	4,676	5,380	4,508	4,430	4,154	3,480	2,978	1,915
10-14 Years (Child)	294	851	1,158	1,700	1,955	1,951	1,996	1,477
15 Years (Young Adult)	1	35	137	379	980	1,672	2,143	3,025
By Race and Ethnicity								
Hispanic	937	1,158	1,304	1,483	1,546	1,520	1,550	1,193
Black	1,604	1,895	1,994	2,133	2,350	2,330	2,229	1,914
Non-Black/Non-Hispanic	2,430	3,213	2,505	2,893	3,193	3,253	3,288	3,310

Major strengths of the NLSY79 child data are the panel and intercohort dimensions of the data collections. It is possible to follow large samples of children across much of their lives. Chapter 4 focuses on analytical connections over the life course, suggesting a number of potential within and cross-disciplinary research possibilities. In this chapter, the thrust of the discussion is on longitudinal dimensions that deal directly with the sample sizes that are available for different kinds of longitudinal research, in a generic sense. This discussion looks at three specific data perspectives: the number of cases available across surveys for children of different ages as of 2000; the possibility for enhancing sample sizes across narrowly defined age groups, by cumulating children of specified ages at different survey points; and the possibilities for sibling and cousin research that exist because the original main respondent sample selection included all individuals in a household that were between the ages of 14 and 21, and the data collection for the children of the female respondents includes multiple interviews with all of the children.

Patterning of Child Interview Frequency

This section provides sample sizes for the number of times younger children and young adults of different ages have been interviewed over the life course of the survey. Given that the child interviewing process began with the 1986 interview round, the maximum number of possible child interviews as of the 2000 survey is eight, beginning with 1986 and incrementing on a biennial basis to 2000. Clearly, the content varies considerably between the child and young adult interviews, partly because in recent rounds only younger children were assessed, and partly because many of the questions are life cycle specific. That is, many questions that might be relevant for an eight- or a fourteen-year old might not be appropriate for an older adolescent. Additionally, there have been some changes in questions and question wording over time that suggest that researchers who are using these data in a longitudinal manner need to carefully review the content of the questions they are using. This issue is considered in various ways in the chapters that follow.

The child and young adult surveys are characterized by inherently different question structure, format of data collection, and indeed potentially different research agendas. It is therefore useful to present separately the sample sizes for younger and older children, even though the ultimate research agenda in many instances may join these two sample types. Also, note that child age in Tables 1.5 and 1.6 references December 31 of each relevant survey year rather than the survey date. This is because the age determination for inclusion as a young adult rather then a younger child was the age as of the end of the calendar year. This method of computing age avoids a need to split the fourteen-year old age group between a younger child and a young adult component. Whether one uses a survey date or end of year age typically has little impact on the magnitude of age-specific sample sizes. Most of the younger child sample size estimates presented in this volume will use child age as of the survey date. This is the reason why sample sizes by age presented across tables may not always be identical. Finally, the ages for the children and the young adults indicated in Tables 1.5 and 1.6 respectively indicate their ages as of the end of 2000, *regardless of whether or not they were interviewed in 2000*. However, the number of interviews reported *is* the actual number of completed interviews between 1986 and 2000 for the children.

Table 1.5. NLSY79 Younger Children: Number of Interviews by Race/Ethnicity and Age as of December 31, 2000

Age of Child	Number of Interviews								
	1	2	3	4	5	6	7	8	Total
All Races									
<1	63								63
1	112								112
2	80	59							139
3	35	162							197
4	28	100	86						214
5	13	55	156						224
6	14	35	127	106					282
7	14	21	79	216					330
8	10	20	52	162	133				377
9	13	15	26	100	252				406
10	17	5	24	68	168	152			434
11	15	16	20	22	125	322			520
12	82	15	15	28	62	161	170		533
13	101	14	25	19	51	114	273		597
14	90	31	16	23	41	101	234	62	598
Total	687	548	626	744	832	850	677	62	5,026
Hispanic Mother									
<1	17								17
1	32								32
2	14	11							25
3	12	28							40
4	8	19	13						40
5	5	18	28						51
6	4	10	27	13					54
7	4	8	23	26					61
8	4	5	19	35	19				82
9	6	4	1	31	43				85
10	7	3	6	20	31	23			90
11	5	2	5	7	37	53			109
12	7	2	4	3	19	39	26		100
13	4	4	6	5	17	37	49		122
14	8	2	7	6	9	35	44	8	119
Total	137	116	139	146	175	187	119	8	1,027

Table 1.5. NLSY79 Younger Children: Number of Interviews by Race/Ethnicity and Age as of December 31, 2000 (continued)

Age of Child	Number of Interviews								
	1	2	3	4	5	6	7	8	Total
Black Mother									
<1	16								16
1	22								22
2	20	12							32
3	10	32							42
4	10	19	18						47
5	2	17	34						53
6	3	10	28	18					59
7	3	7	35	53					98
8	2	1	14	39	30				86
9	2	4	9	35	49				99
10	5		4	23	42	34			108
11	3	4	7	5	48	90			157
12	4	3	1	9	22	50	48		137
13	2	5	6	2	12	38	72		137
14	5	5	4	6	15	39	65	19	158
Total	109	119	160	190	218	251	185	19	1251
White Mother									
<1	30								30
1	58								58
2	46	36							82
3	13	102							115
4	10	62	55						127
5	6	20	94						120
6	7	15	72	75					169
7	7	6	21	137					171
8	4	14	19	88	84				209
9	5	7	16	34	160				222
10	5	2	14	25	95	95			236
11	7	10	8	10	40	179			254
12	71	10	10	16	21	72	96		296
13	95	5	13	12	22	39	152		338
14	77	24	5	11	17	27	125	35	321
Total	441	313	327	408	439	412	373	35	2748

Table 1.6. NLSY79 Young Adult Children: Number of Interviews by Age and Race/Ethnicity

Age of Young Adult	# of Young Adult Interviews				# of Child Interviews			Total # of Interviews			
	1	2	3	4	0-5	6	7	0-5	6	7	8
All Races											
15	383				33	40	310	11	22	40	310
16	364				31	35	298	17	14	35	298
17	118	327			110	335		26	46	106	267
18	126	320	1		92	355		30	25	120	272
19	37	125	262		424			32	39	126	227
20	40	113	247		400			37	45	106	212
21	21	19	53	238	331			35	25	59	212
22	26	14	53	149	242			38	21	59	124
23–25	54	94	225	89	462			100	92	195	75
26–30	17	12	53		82			27	27	28	
Total	1,186	1,024	894	476	2,207	765	608	353	356	874	1,997
Hispanic Mother											
15	74				8	10	56	2	6	10	56
16	69				6	14	49	4	2	14	49
17	38	73			28	83		5	11	38	57
18	39	67			39	67		11	12	35	48
19	8	45	53		106			12	9	40	45
20	8	31	47		86			11	13	25	37
21	8	9	12	50	79			15	8	12	44
22	9	5	8	35	57			14	5	14	24
23–25	6	20	44	21	91			18	23	35	15
26–30	4	3	6		13			6	3	4	
Total	263	253	170	106	513	174	105	98	92	227	375
Black Mother											
15	107				11	15	81	5	6	15	81
16	98				7	6	85	4	3	6	85
17	57	104			47	114		9	27	43	82
18	62	98	1		27	134		7	8	65	81
19	10	53	91		154			9	16	49	80
20	12	57	83		152			12	15	50	75
21	5	6	25	107	143			10	12	31	90
22	9	2	31	70	112			14	8	30	60
23 to 25	28	45	115	42	230			53	39	100	38
26 to 30	9	6	38		53			14	20	19	
Total	397	371	384	219	936	269	166	137	154	408	672
White Mother											
15	202				14	15	173	4	10	15	173
16	197				18	15	164	9	9	15	164
17	23	150			35	138		12	8	25	128
18	25	155			26	154		12	5	20	143
19	19	27	118		164			11	14	37	102
20	20	25	117		162			14	17	31	100
21	8	4	16	81	109			10	5	16	78
22	8	7	14	44	73			10	8	15	40
23 to 25	20	29	66	26	141			29	30	60	22
26 to 30			9		16			7	4		
Total	526	400	340	151	758	322	337	118	110	239	950

Table 1.5 indicates the total number of interviews reported for each of the younger children who have been interviewed at some point since 1986. We use 1986 as the starting point for this cumulative count, as 1986 was the first year that children of the NLSY79 were assessed. As noted above, the greatest number of possible interviews for any child in 2000 would be eight biennial interviews between 1986 and 2000. However, to have eight interviews, a child would have had to have been a newborn in 1986, have been born early in the year (most 1986 interviews occurred early in that year), and have been interviewed at all eight possible survey points. As may be seen, only 62 of the younger children age 14 at the end of 2000 fall into that category. However, Table 1.5 shows that much larger numbers of children fall into all the other interview frequency categories. Children who fall into the older age categories as of 2000, but who have completed only a small number of interviews (e.g., 11 year olds with only one or two interview points) have missed some interviews. For example, an 11 year old in 2000 should have six, or at least five completed interviews. Note that there are 73 eleven year olds as of 2000 who have completed four or fewer interviews. The implications of interview repetition for these younger children will be expanded on in Table 2, where this interviewing repetition will be connected with the specific cognitive and socio-emotional assessments that the children complete at younger and older ages.

Table 1.6 extends this interview repetition notion to the young adult component of the survey. Since the young adult survey, which is detailed in Chapter 3, has been ongoing only since 1994, the maximum number of young adult interviews possible by 2000 is four. This maximum would be limited to young adults who are at least 20 years of age as of the end of the 2000 calendar year. It should be recalled that young adults age 21 or over at the end of the 1998 interview year were not interviewed in that year. This leads to a four-year interview gap, between 1996 and 2000, for young adults who are 23 or older as of the end of 2000. From year 2002 forward, the expectation is that the young adult sample will include all youth age 15 and over. While the age references the end of 2000, a particular young adult was not necessarily interviewed in that year. Of the 3,580 young adults included in the table, 3,025 (about 84 percent) were interviewed in 2000. Table 1.6 shows that there are substantial numbers of young adults who have four young adult interview points and a very large sample who have had three interviews as young adults since 1994. Cumulatively, across both the child and young adult interviews, there are almost 2000 young adults who have been

interviewed every round since 1986. In this regard, there are relatively large samples of black, Hispanic and non-Hispanic white (hereafter termed white) young adults who fall in this eight-interview category.

Table 1.6 shows that most of the young adults have had a number of interviews prior to reaching age 15, and then between one and four interviews since that date. Chapter 2 will detail the kinds of information that was collected from and about the younger children, and Chapter 3 will focus on the data collection since age 15. The nature of the data collection changes in fundamental ways when a child transitions from being part of the younger cohort to being a young adult. Below age 15, children are administered (or their mother completes) a variety of assessments. At all these ages the mother also provides a variety of information about the child's health, education, and selected other items, all of which are detailed in Chapter 2. Since 1988, children age ten and over have completed questions about their own behaviors and attitudes in a variety of domains, including education, family and peer interaction processes, selected normative and non-normative attitudes and behaviors and other linked activities. Since 1994, once attaining age 15 (subject to the sampling caveats addressed above or in the chapters that follow), the young adults have completed a lengthy personal interview that addresses issues that encompass most of the major dimensions of their lives—schooling, employment, family, peer interactions and issues of sexuality, and a variety of other behaviors and attitudes that permit the examination in context of their experiences in a holistic manner. A flow diagram that would cross the possible eight survey points for these youth encompassing the 1986 to 2000 period would describe a coherent data collection process that typically follows a child from his or her preadolescent years, describing the child development process in some detail, parallel family and child behaviors and attitudes, and culminate in a kaleidoscope, or moving picture of his later adolescent-early adult transitions to adulthood. The younger child and young adult components of this process are highlighted in Chapters 2 and 3 respectively. Chapter 4 attempts a synthesis of selected data elements and suggests some research agendas for which this data set is particularly appropriate.

Pooling Sample Sizes

The panel dimension of the NLSY79 data collection permits one to cumulate sample cases for children at specified ages across survey points, thus attaining rather substantial sample sizes for children at specific ages. Pooling in this manner also can greatly enhance the

heterogeneity of the sample for specific research topics. The trade off to this methodology is that the ability to follow a particular age cohort across survey years becomes somewhat more limited, although it is still doable for selected research topics. Tables 1.7 and 1.8 highlight potential sample sizes using this approach.

Table 1.7. NLSY79 Younger Children: Sample Sizes for Pooled Age Groups across 1986-2000 Survey Years

Pooled Age Groups	Total Possible #
0 year olds	2,641
1 year olds	2,944
2 year olds	3,082
3 year olds	3,267
4 year olds	3,298
5 year olds	3,384
6 year olds	3,326
7 year olds	3,335
8 year olds	3,194
9 year olds	3,047
10 year olds	2,920
11 year olds	2,660
12 year olds	2,402
13 year olds	2,157
14 year olds	1,323

NOTE: Ages at survey dates: Cumulative from 1986 to 2000 interview. Children interviewed at least at one survey point.

Table 1.8. NLSY79 Young Adult Children: Sample Sizes for Pooled Age Groups across 1994-2000 Survey Years

Pooled Age Groups	Total Possible #	# of Males	# of Females
14 Year Olds	779	381	398
15 Year Olds	1,450	734	716
16 Year Olds	1,292	648	644
17 Year Olds	1,068	549	519
18 Year Olds	970	463	507
19 Year Olds	685	353	332
20 Year Olds	625	298	327
21 Year Olds	329	178	151
22 Year Olds	218	102	116
23 Year Olds	155	74	81
24 Year Olds	126	67	62
25 Year Olds	67	34	33
26 Year Olds	29	13	16
27 Year Olds	18	10	8
28 Year Olds	3	1	2
29 Year Olds	1	0	1

NOTE: Ages at survey dates: cumulative from 1994 to 2000 interview. Young Adults interviewed in at least one year.

From the perspective of the younger children, it may be seen that it is possible to attain single year of age samples numbering in the thousands for specific research efforts. For example, if one wishes to examine associations between memory for digit span scores and other factors for seven year olds, it would be possible to cumulate a sample of more than 3,000 seven year old children across the 1986 to 2000 survey years. Because these children would have been born to mothers in all years between 1979 and 1993 (see Table 1.3), the mothers of these children would range in age from their early teens to their early thirties. The relevance of this pooling approach for younger child evaluation, utilizing various assessments, will be considered further in Chapter 2.

Parallel estimates are provided in Table 1.8 for young adults. In this instance, while several of the cumulated sample sizes are substantial, they do not attain the magnitudes of the sample sizes for the younger children. This is because the young adult interviews have only been on-going since the 1994 survey round, so no single year of age cumulative estimate can include more then four points. Nonetheless, the number of cases cumulated in this way for ages 15 to 18 all attain almost 1000 cases or higher. Also evident in Table 1.3 is that this sample cumulation modestly increases the heterogeneity of these young adult samples. For example, cumulating cases at these early young adult ages expands the age of the mothers at

children's birth from mid-adolescence to the late twenties. Of course, using appropriate statistical technology permits one to collapse sample cases across several young adult age groups, building to a very large, quite heterogeneous sample size.

Sibling and Cousin Samples

When the sample selection for NLSY79 was made, all individuals living in the selected households who were between the ages of 14 and 21 on December 31, 1978 were selected for sample inclusion. In many instances, siblings were included in the original sample. This has methodological implications for those who are concerned about the lack of complete independence between all of the NLSY79 cases. From the perspective of the children of the NLSY79, the particular focus here is on the main Youth sisters who are respondents in the sample. The *NLSY79 User's Guide* details this information, so it is not considered extensively here. The focus in this section is on the children who have been born to the female respondents in this sample. From the child's perspective, children of sisters are cousins to each other. Over the course of the survey years, more than 3000 children in the sample have been identified as having an aunt in the main NLSY79 sample. Most of these children have one aunt, but smaller numbers have multiple aunts. While the number of children who are cousins is considerable, the precise numbers available for a particular research project are contingent on the objectives of the research; in particular, will the researcher be limiting his/her sample to children or women interviewed in 2000 or will the researcher include mothers or children interviewed in one or more earlier survey rounds?

More typically, researchers have been utilizing the large number of *child* sibling sample cases that have been born to the female respondents. As seen in Table 1.9 (which focuses on the children of women interviewed in 2000), most of the women interviewed have had more then one child, including a rather large sample of women who have had three or more children, as they approach the end of their childbearing years. In addition to multiple births, there are many family units where the two or more children are widely spaced in age, thus enhancing the possibility of exploring the impact of childbearing on children that have been born to the same mother but at different maternal life cycle stages. This table is limited to women who have been interviewed in 2000 and their children. Larger sample sizes can be generated by incorporating into one's sample women who were not interviewed in 2000 but who had been interviewed in earlier survey rounds.

Table 1.9. NLSY79 Women Interviewed in 2000 by Number and Ages of Children and by Race/Ethnicity

Type of Household (Female)	Age of Child(ren)	Number of Households			
		Hispanic	Black	White	Total
Females with no child		100	205	383	688
Mothers with 1 child	< 6 years old	24	21	74	119
	6-9 years old	14	31	73	118
	10-14 years old	30	46	91	167
	15+ years old	40	113	124	277
	Total	**108**	**211**	**362**	**681**
Mothers with 2 children	Both < 6 years old	11	14	49	74
	Both 6-9 years old	8	10	45	63
	Both 10-14 years old	19	24	67	110
	Both 15+ years old	69	129	191	389
	Older 6-9, Younger < 6	17	15	76	108
	Older 10-14, Younger < 6	14	11	25	50
	Older 15+, Younger < 6	10	15	10	35
	Older 10-14, Younger 6-9	29	34	120	183
	Older 15+, Younger 6-9	15	31	24	70
	Older 15+, Younger 10-14	67	86	167	320
	Total	**259**	**369**	**774**	**1,402**
Mothers with 3 or more children	All < 6 years old	2	1	5	8
	All 6-9 years old	0	1	1	2
	All 10-14 years old	2	2	4	8
	All 15+ years old	50	113	69	232
	Oldest 6-9, Youngest < 6	8	3	40	51
	Oldest 10-14, Youngest < 6	16	20	62	98
	Oldest 15+, Youngest < 6	65	74	71	210
	Oldest 10-14, Youngest 6-9	14	17	47	78
	Oldest 15+, Youngest 6-9	67	92	102	261
	Oldest 15+, Youngest 10-14	93	150	151	394
	Total	**317**	**473**	**552**	**1,342**

As detailed elsewhere in this report, sisters, as well as children born to those sisters can be readily identified. NLSY79 female respondents who are sisters and who were resident in the same household when the sample was selected can be identified by variables on the child file called SISTID1-3 (C00010.-C00012.)**.** Children born to a particular respondent all share the same stem as the ID of their mother, with an additional two-digit identifier (01, 02 etc.) that typically (although not in all cases) clarifies their sibling placement.

Using the Sampling Weights

The 2000 sampling weights for younger children (C24955., CSAMWT2000) and young adults (Y11923., YA00WEIGHT) (1) adjust the unweighted data for sample attrition of mothers and their children since the first survey round (1979) and the sample reduction due to the loss of the military and economically disadvantaged white oversample and (2) adjust the

sample for the over-representation of black and Hispanic youth. For those interested in generating population estimates for prior survey rounds, sampling weights for those survey rounds are available. Using these weights translates the unweighted sample of children into a population that represents all children who have been born by that date to a nationally representative sample of women who were 14 to 21 on December 31,1978.

Weights are computed only for younger children or young adults who have been interviewed in a given year. *Children not assessed and young adults not interviewed in a given year are assigned a weight of zero for that year*. The child's sampling weight equals the mother's 1979 weight (C00700.04, SAMPWT79) times an adjustment factor that is the reciprocal of the rate at which children in a particular age, sample type, and sex cell are interviewed. When these individual cells contain small counts, they are grouped across ages (and sometimes sample types) to avoid large fluctuations in the child weights. This grouping varies from year to year, primarily because of the increasing sample sizes over time at the higher young adult ages. Weights can be particularly unstable at the oldest ages where sample sizes may be small. Table 1.10 lists the complete set of child, young adult, and mother sampling weights.

Table 1.10. NLSY79 Child, Young Adult, and Mother Sampling Weights

Reference Number	Variable Description	Question Number	Year	Area of Interest
C00700.04	Sampling Weight of Mother	SAMPWT79	1979	FAMILY BACKGROUND
C05812.	Child Sampling Weight	CSAMWT86	1986	ASSESSMENT 1986
C08007.	Child Sampling Weight	CSAMWT88	1988	ASSESSMENT 1988
C09999.	Child Sampling Weight	CSAMWT90	1990	ASSESSMENT 1990
C11999.	Child Sampling Weight	CSAMWT92	1992	ASSESSMENT 1992
C15089.	Child Sampling Weight	CSAMWT94	1994	ASSESSMENT 1994
C15658.	Child Sampling Weight	CSAMWT96	1996	ASSESSMENT 1996
C18012.	Child Sampling Weight	CSAMWT98	1998	ASSESSMENT 1998
C24955.	Child Sampling Weight	CSAMWT2000	2000	ASSESSMENT 2000
Y03565.	Young Adult Sampling Weight	YA94WEIGHT	1994	YA COMMON KEY VARIABLES
Y06507.	Young Adult Sampling Weight	YA96WEIGHT	1996	YA COMMON KEY VARIABLES
Y09469.	Young Adult Sampling Weight	YA98WEIGHT	1998	YA COMMON KEY VARIABLES
Y11923.	Young Adult Sampling Weight	YA00WEIGHT	2000	YA COMMON KEY VARIABLES

We caution users that comparing weighted estimates across years can be risky as the composition of the sample can change in subtle ways depending on who was interviewed.

Analyzing data from children interviewed in multiple years also creates problems since the yearly weights are not appropriate to such a universe. To be correct, weights for a multi-wave analysis would have to be constructed for that particular set of observations. If analyses are limited to children present in one specific outcome year, that year's weights should be used. For multivariate analyses utilizing samples across survey rounds, using the unweighted data is encouraged.

Research Based on the NLSY79 Child and Young Adult Data

The within and cross generation research possibilities offered by this data set are considerable. They overlap the interests of researchers in a rather wide range of intellectual disciplines. This chapter has highlighted a number of data and sample considerations that are relevant to many research agendas, suggesting strengths and limitations of this panel data set. Chapters 2 and 3 provide considerable detail about the contents of the younger child and young adult data files. Chapter 4 suggests how the data set can be used for mainstream research topics in sociology, economics, child development, public health, and several other related disciplines.